DOCUMENT RESUME

ED 173 211 SO 011 769

AUTHOR Carroll, Glenn R.; And Others
TITLE Final Report for Dynamic Models for Causal Analysis of Panel Data. The Impact of Measurement Error in the Analysis of Log-Linear Rate Models: Monte Carlo Findings. Part III, Chapter 4.
INSTITUTION Stanford Univ., Calif.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Nov 78
GRANT NIE-G-76-0082
NOTE 49p.; For related documents, see SO 011 759-772

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data Analysis; *Educational Change; Measurement Techniques; Models; *Research Methodology; *Research Problems; *Social Change; Social Science Research; Sociometric Techniques; *Validity

ABSTRACT

This document is part of a series of chapters described in SO 011 759. The chapter advocates the analysis of event-histories (data giving the number, timing, and sequence of changes in a categorical dependent variable) with maximum likelihood estimators (MLE) applied to log-linear rate models. Results from a Monte Carlo investigation of the impact of measurement error on the performance of MLE applied to a log-linear rate model are reported. Sections of the document describe constant rate models of both a single event and a multiple event, examine complications created by errors in recording the timing of events, focus on the same complication for the multivariate rate model, and consider consequences of measurement error in the exogenous variables for this model. Results indicate that the performance of MLE in rate models is good. Also, measurement error in the waiting time and in exogenuous variables causes only minimal deterioration in estimator quality. The studies suggest that the most pronounced effects of measurement error can be avoided if the investigator does not impose arbitrary censoring points upon complete event-histories. When data on the timing of natural events are available, they should be used in entirety. (Author/KC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173211

AFR 2133 50

PART III - Chapter 4

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report for

DYNAMIC MODELS FOR CAUSAL ANALYSIS OF PANEL DATA

THE IMPACT OF MEASUREMENT ERROR IN THE ANALYSIS OF LOG-LINEAR RATE MODELS: MONTE CARLO FINDINGS*

Glenn R. Carroll
Michael T. Hannan
Nancy Brandon Tuma
Barbara Warsavage

November 1978

SE 011 769

*This research was supported by National Institute of Education grant NIE-G-76-0082. This report is also available as Technical Report #69 of the Laboratory for Social Research, Stanford University.

1. Introduction

Recent methodological work in sociology has devoted much attention to the modeling of discrete events, especially over time.[1] In our research we have advocated the analysis of event-histories with maximum likelihood estimators (MLE) applied to log-linear rate models (Hannan, Tuma and Groeneveld 1977; Tuma and Hannan 1978; Tuma et al. 1979). We recommend MLE primarily because of its asymptotic properties: consistency, efficiency and normality under fairly weak regularity conditions on the probability distribution function specified by the model (see Dhrymes 1970). Other advantages of MLE include the ease with which censoring can be treated and the minimal effort required to generalize the procedure to a wide range of models (see Tuma and Hannan 1978). Nonetheless, we recognize that sociologists do not have much experience with these methods; many properties of MLE are not well understood, especially in comparison with other structural models used in sociology. For this reason, we have embarked on a research program to expand knowledge in this area. After briefly reviewing this program, we report the results from a Monte Carlo investigation of the impact of measurement error on the performance of MLE applied to log-linear rate models.

2. Background

Our research builds on the work of Bartholomew (1957, 1959, 1963) and Mendenhall and Lehman (1960). In investigations of models in which the rate of an event or hazard is assumed to be the same for all individuals, these investigators found ML estimates to be slightly upward biased in

small samples. A Monte Carlo study by Tuma and Hannan (1978) replicated these findings and generalized them to research contexts with censored observations and multiple events.[2] Other work in our program has concentrated on the MLE of models with population heterogeneity. These models relax the assumption that the rate is constant within the population and allow each individual's rate to depend on specific exogenous characteristics. In the first of these investigations, Keeley (1975) studied MLE in a model in which the rate is a log-linear function of uniformally-distributed independent variables. He found that MLE gave estimates of good quality, even when the vector of independent variables is misspecified. In contrast, Fennell, Tuma and Hannan (1977) obtained poor quality MLE under misspecification of a similar log-linear rate model with normally-distributed exogenous variables. Still, they also found that MLE of the correctly specified model were consistently high in quality, although slightly upward biased, under varying conditions of censoring and of collinearity between the independent variables.

In this study, we report the findings from several Monte Carlo experiments designed to extend these results when there is measurement error. We begin with a model in which there is a constant rate of a single event and then move to a model with multiple events. For both experiments, we consider complications created by errors in recording the timing of events. The final section of this report focuses on the same complication for the multivariate rate model, i.e., observable population heterogeneity. We also consider the consequences of measurement error in the exogenous variables for this model. Throughout the paper we concentrate on practical research problems, focusing especially on the

interaction between censoring and measurement error.

3. A Constant-Rate Model of a Single Event

We begin with the continuous-time, discrete-state Markov process discussed at length by Tuma and Hannan (1978). We let j and k represent any pair of states in the system, t represent a point in continuous time, Δt represent a nonnegative time interval, and $Q_{jk}(t,t+\Delta t)$ represent the probability of a change from state j at time t to state k at time $t+\Delta t$. $Q_{jk}(t,t+\Delta t)$ is usually termed a transition probability and differs from $q_j(t)$, the probability of being in state j at time t, which is sometimes termed the state probability.

We define $r_{jk}(t)$, the rate of change from j to k at time t, as:

$$r_{jk}(t) = \lim_{\Delta t \to 0} \frac{Q_{jk}(t,t+\Delta t)}{\Delta t} \quad , \quad j \neq k \tag{1}$$

and $r_j(t)$, the rate of leaving j at time t, as:

$$r_j(t) = \lim_{\Delta t \to 0} \frac{1-Q_{jj}(t,t+\Delta)}{\Delta t} \tag{2}$$

Further, we assume that $r_{jk}(t)$ exists for every pair of states j and k, $j \neq k$, that $r_{jk}(t)$ is a continuous function of time t, and that $0 \leq r_{jk}(t) < \infty$, for all j, k and t.

We let Y(t) record the number of events that occur by time t, $0 \leq t < \infty$, and let the rates that define the Markov process be given by

$$r_{jk} = \begin{cases} \alpha & \text{if } k = j + 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Thus transitions can only occur to the next highest integer, i.e. Y(t) is a nondecreasing function of time and in any small moment of time increases by at most one. This model is simply the well-known Poisson model (see Breiman 1969 for a lucid discussion). In particular:

$$\Pr[Y(t) = y] = \frac{(\alpha t)^y}{y!} e^{-\alpha t} \tag{4}$$

where t is the length of the observation period.

We are interested in estimating α, the rate at which events occur, for samples both with and without censored observations. To accomplish this, we use the following ML estimator (see Tuma and Hannan 1978):

$$\alpha_0 = \sum_{i=1}^{N} \frac{Y_i}{Y_i u_i + (1-Y_i)\tau_i} \tag{5}$$

where Y_i is one if the event occurs before τ_i for individual i and zero otherwise, and u_i is the waiting time until the event for individual t.

In the Monte Carlo study we set α to equal 1 per unit of time. This yields a model with both mean duration between events and variance equal to 1. To simulate u, the waiting time until an event for each sample unit, we used the relationship

$$\ln [G(u)] = \ln [e^{-\alpha u}] = -\alpha u = -u \tag{6}$$

where G(u) is a pseudo-random uniform deviate. That is, for each unit we drew a pseudo-random number from a (0,1) uniform distribution and called this G(u); we then used (4) to calculate u. We repeated this

entire process 100 times for each sample size of 25, 50 and 100 units.

In this report we study two different types of measurement error, random and time-dependent. Measurement error is simulated by altering the times generated by (4). Specifically, we calculate

$$t^{*\ell} = \gamma_\ell u \tag{7}$$

where γ_1 represents a pseudo-random deviate from a gamma distribution with mean 1 and variance one-tenth and γ_2 denotes a similar deviate but with variance one-tenth of the waiting time u.[3] These changes do not alter the expected value of u but do modify its variance. For the random measurement error experiment, the variance of t^{*1} is 1.1 while for the time-dependent experiment the variance of t^{*2} equals 1.6. Thus, if we measure reliability as $R = \frac{\text{var}(u)}{(t^{*\ell})}$, then our simulation represents research settings of event history data with .91 reliability (random experiment) and .63 reliability (time-dependent experiment) in the observed value of the waiting time.

One additional issue must be addressed before turning to the results: the interaction effects of censoring and measurement error in the time variable. In our experiments, we determined the censoring times (τ) from expected proportions (.10, .20, .50, .80, 1.0) generated by the underlying model. We did not alter the censoring points to compenstate for changes in the distribution of the waiting time caused by the disturbances. Thus, alteration of a waiting time because of measurement error sometimes shifts the timing of an individual's event across a previously established censoring boundary value.[4] We call this phenomena <u>false censoring</u> since

the intraction of measurement error and censoring shifts the observation of events from the true period to another.

Problems of false censoring are likely to arise when the investigator imposes censoring levels upon event-history data. For example, a sociologist studying merger activity in the American economy might confine the analysis to a certain historical period. For the sample of organizations followed during this period, certain ones will not be involved in mergers and hence will be censored observations. However, some of these censored cases may have actually been involved in a merger during the period of concern but the date of the merger erroneously recorded as occurring before or after the observation period. In this case, measurement error creates false censoring. If the entire event history is used in the analysis, measurement error alters the value of the time variable but does not affect the actual number of events. By imposing an observation period, the investigator creates an additional error in the number of events observed in the period; this is false censoring.[5]

We have not yet discussed the consequences of false censoring for MLE based on event-history data. Frankly, we are uncertain what to expect. For this reason, we pay particular attention to this issue in the results discussed below.

Results. We report the experimental outcomes for the single event model in Table 1. For each sample size and censoring level we display the mean and variance of the estimator over the 100 samples. When there is no measurement error in the waiting time, the ML estimator is slightly upward biased for all sample sizes. Moreover, bias is apparently unaffected

by the level of censoring. This contrasts with the behavior of the estimator's variance, which increases substantially with censoring. These findings agree with those of previous investigators (Mendenhall and Lehman 1960; Tuma and Hannan 1978). For clarity, we have summarized them by plotting the distribution for the ML estimator at various censoring levels in Figure 1. The curves in this figure show that at high levels of censoring the ML estimator is not only more dispersed, but also less symmetric. This asymmetry would eventually disappear if we increased the number of trials since the estimator is theoretically normally-distributed at all censoring levels.

Table 1 also shows the estimator's performance when dates are recorded with error. As we expected, measurement error damages the quality of MLE, both the bias and the variance increase, although not dramatically. Several additional patterns are noteworthy. For both types of disturbance the upward bias increases. This pattern at least partially reflects the parametric form used to simulate the disturbances. Although the gamma distribution used in both cases has a mean of 1, the median is slightly less than 1. Consequently, the simulated disturbance lowers more waiting times than it raises. Thus, we speculate that this process causes the estimated rate to be higher than the true rate.

The other interesting result in Table 1 concerns the effect of censoring when there is measurement error in the waiting time. For random error we notice a distinct pattern; bias increases with censoring. This pattern was not observed, however, in the case of accurate measurement. Thus we believe the pattern of increasing bias is caused by false

censoring. In the data with a time-dependent disturbance, the same effect occurs, but in a less obvious manner. The troublesome finding is that the bias decreases for the highest censoring levels. This drop occurs because events observed at the 90% censoring level have very small waiting times, thereby causing their time-dependent disturbances to vary only a little. Consequently, effects due to time-dependent measurement error are minimal at high levels of censoring. This contrasts with the random disturbance-case, where at 90% censoring, the variance of the disturbance (and thus its impact) is greatest relative to the waiting time. Therefore, we conclude that the decrease in bias at high levels of censoring when there is time-dependent error results from a lower level of false censoring than in medium levels of censoring, i.e., 50%.

These findings paint a dismal picture of the consequences of investigator-imposed censoring of complete event-histories. For as we have seen with this model, such censoring runs the risk of creating false censoring when errors confound the data. Figure 2 shows the impact of this problem on MLE in our Monte Carlo studies. Measurement error causes the bias of the estimator to shift significantly upward, and the distribution curve to be more widely dispersed. Certainly this represents a situtation to be avoided if at all possible.

4. Multiple Event Constant-Rate Model

Sociologists often study types of events which recur during an individual's lifetime, e.g. marriages, divorces, and job changes. Using the procedure described in Tuma and Hannan (1978), we are able to simulate

such multiple event data.[6] We again use equation (4) to simulate waiting times, except we index j events for each individual i until the total time observed ($\Sigma_j u_{ij}$) exceeds five units (which we now assume to be years for convenience). We then censor observations on each individual at five. Since the expected rate and time still equal 1, the expected number of events for each individual during the five-year period is five. Of course, some units have more and others fewer.

We examine the ML estimator under several different conditions. The more familiar application is a generalized version of equation (5), of the form

$$\hat{\alpha} = \frac{Y}{U+V} \tag{8}$$

where Y is the total number of events for all individuals occurring during the observation period, U is the total time spent by individuals in completed spells and V is the total time spent in censored spells. The other estimator used is the so-called backward MLE described by Tuma and Hannan (1978):

$$\hat{\alpha}_B = \frac{1-\frac{Y_0}{N}}{\overline{W}} \tag{9}$$

where Y_0 is the number of individuals with no events during the observation period, $\overline{W}$ is the mean length of censored spells and N is the sample size. This estimator uses information only on the length of the last spell when the process is near the origin and is useful when the starting dates of the last spell are unknown.

We study the quality of these estimators under varying conditions of investigator-imposed censoring. Initially we use data for the first year

only, a situation in which ensoring is imposed at the end of the observation period. We estimate both MLE, an application of (8) defined as $\hat{\alpha}_1$, and backward MLE, which uses (9) and is called $\hat{\alpha}_2$. Next we examine estimators where censoring is imposed at the beginning of the fifth year. Again we examine MLE, which uses (8) for this observation period and is labelled $\hat{\alpha}_4$, and backward MLE, which is estimated with (9) and is defined as $\hat{\alpha}_5$. For comparison, we estimate an all-event MLE ($\hat{\alpha}_3$) which applies (8) to data for the entire period. We also examine two ML estimators for data based on the timing of the first event. The first of these, $\hat{\alpha}_6$, uses (8) where the relevant observation period is the duration until the first event of five and thus varies by individual. The second estimator, $\hat{\alpha}_7$, also uses (8) but in this case the observation period is the period between the first spell and the end of the entire process observed, in this case five. Again, the relevant observation period for this estimator varies by individual.

These experiments again incorporate disturbances as in equation (7). However, the variance of the time-dependent disturbance now depends on the length of time away from the end of the observation period ($5-t_j$) rather than the duration of the spell.[7] Thus, these experiments represent a situation in which recall is less accurate for events further from the end of the observation period. Such a phenomena is likely to arise when subjects are asked to give retrospective event-histories.

Results. Table 2 presents the results for this portion of the study. When there is no measurement error, each estimator performs well, except for the backward MLE for the fifth year ($\hat{\alpha}_5$). The simulated disturbances produce some confusing results. For the first year estimators ($\hat{\alpha}_1$ and $\hat{\alpha}_2$),

measurement error increases bias. As before, investigator-imposed censoring apparently exacerbates the deleterious effects of measurement error. However, the results for the doubly-censored fifth year estimators ($\hat{\alpha}_4$ and $\hat{\alpha}_5$) are less clear-cut. The case of time-dependent disturbances performs as expected but the case of random disturbances results in lower, rather than higher, estimates of the rate. A clue to explaining these findings can be found again in the parametric form used for the disturbances. The gamma disturbance more frequently lowers waiting times and thus increases estimates. However, for a given five-year observation period the number of events is fixed. Thus, the gamma disturbance shifts events to an earlier point in time. Finding a lower estimated rate in the fifth year is not surprising then, since events have been disproportionally shifted to an earlier period. In the time-dependent case, this shift is less pronounced. Because the disturbances are smaller in the fifth year (recent events are more accurately recalled), the events are not shifted as much as in the first year. However, we should still expect some events to be shifted out of the fifth year and consequently estimate a lower rate. The higher estimated rate of $\hat{\alpha}_4$ for the time-dependent data is an apparent anamoly.[8]

The performance of the estimators based on the timing of the first event ($\hat{\alpha}_6$ and $\hat{\alpha}_7$) is more easily understood. The first of these, $\hat{\alpha}_6$, estimated an increased rate because the disturbances decrease the duration of the first spell. The other estimator, $\hat{\alpha}_7$, must then estimate a decreased rate because while the number of events remains constant, the observation period ($5-t_{11}$) has more frequently been increased.

If we again consider the issue of investigator-imposed censoring,

several items deserve mention. Foremost, the change in the direction of the bias due to the combination of censoring and measurement error can be predicted less clearly when there are data on multiple events. However, an increase in the absolute value of the bias is consistently observed in this situation. The exception to these remarks is the ML estimator using data over the entire five-year period ($\hat{\alpha}_3$); it is completely insensitive to measurement error.

4. Multivariate Time-Independent Rate Model

Sociologists have historically shied away from Markov models similar to those discussed above because their use implies an unrealistic assumption of population homogeneity. Multivariate rate models overcome this obstacle. We use a particular form of this type model discussed at length by Fennell, Tuma and Hannan (1977). Specifically, we designate the rate to be

$$\ln(r_{jk}) = \begin{cases} \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 & \text{if } k = j + 1 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where X_1 and X_2 are normally-distributed exogenous variables. For this model, the ML estimator does not have a closed-form solution. Instead, we maximize the likelihood function with the Newtonian iterative technique developed by Gill and Murray (1972) and implemented by Tuma and Crockford (1976). In this simulation, we replicate the procedures that Fennell, Tuma and Hannan (1977) used to generate and estimate 100 samples of data.

Replication. Our first effort with this model does not involve

measurement error. Rather, we wanted to replicate and extend the results of Fennell, Tuma and Hannan (1977). These investigators failed to observe a consistent directional pattern of the bias for the model in (8). We suspect these results are not general and might depend on the specific parameters employed. For that reason, we repeated their study using different parameter values. We chose two sets. In the first set we set $\alpha_0 = 4$, $\alpha_1=1$, and $\alpha_2=2$. In the second, $\alpha_0=-4$, $\alpha_1=-1$ and $\alpha_2=1$. For each set of parameters we generated and studied three levels of collinearity between the X's ($\rho=0.0$, 0.5 and -0.5) and three levels of censoring (uncensored, 60% censored and 80% censored). We studied sample sizes of 50 and 100.

Results. In Tables 3 through 6, we report the results for the Monte Carlo study of this model. Table 3 reports the bias of the estimates and Table 4 presents the percent-bias. These findings are consistent with those of Fennell, Tuma and Hannan (1977); ML estimators of this model have only a small bias even at high levels of censoring. However, these results also display a pattern not found in their study. Table 3 shows that the direction of the bias changes as censoring increases. When the parameter is negative, the bias for uncensored data is usually positive. As censoring increases, the bias decreases and becomes negative. The reverse is true for positive parameters. They are usually negatively biased for uncensored data and become positively biased with increased censoring. While this pattern reappears consistently, we know of no theoretical explanation of its occurrence, and thus caution against its generalization.

Table 5 presents data on the efficiency of MLE for this model. Again, our findings replicate the extend those of Fennell, Tuma and Hannan

(1977). We too find low variability in the ML estimator for this model even at high censoring levels. The extension concerns the general decrease in estimator variance for the parameter set of -4, -1 and 1. This decreased variance can best be explained by the lower expected times and rates of the model for this parameter set.[9] Thus, we would expect even lower estimator variance for certain parameter values.

We summarize these findings of the bias and efficiency of MLE for this model in Table 6 and Figure 3. Table 6 presents the mean-squared error (MSE) of the estimates and Figure 3 visually displays the effect of censoring on the quality of the estimates. Table 6 shows that the MSE is extremely low for all conditions, although it increases with censoring. Figure 3 shows this deterioration in the quality of estimates through plots of the estimators' distribution. As shown in this figure, high censoring changes the form, the mode and the dispersion of estimates. However, even at high censoring levels, MLE provides reasonably good estimates.

Measurement Error in the Waiting Time. At this point we again consider measurement error in the waiting time. Using the parameter set -4, -1 and 1, we simulate random measurement error as in equation (5). For this experiment, however, we use a gamma-disturbance with mean 1 and variance ½. This disturbance does not change the expected times until an event but does increase the variance of the waiting times by 50% when the intercorrelation between the exogenous variables is zero.[10] Consequently, the incorporation of simulated measurement leads to a reliability of 1.5 on the waiting time. As in the previous studies, we again examine three censoring levels and allow measurement error to create false censoring.

Results. Tables 7 through 10 present the findings from this experiment as contrasted with those without measurement error. Tables 7 and 8 report the bias and %-bias, respectively. In Table 7, we observe a new trend. Without censoring, the measurement error findings closely resemble the results when there is no measurement error, although the bias is slightly larger. With censoring this pattern changes. ML estimators for the data with error at this level tend to overestimate negative parameters (have an upward bias) and to underestimate positive parameters (have a downward bias). In other words, investigator-imposed censoring moves the parameter estimates for this model closer to zero. Thus, if the investigator were performing causal analysis, estimates of the association between the exogenous variables and the rate of events would be weaker than actually exists. This is exactly opposite the pattern for censoring bias when there is no measurement error. We remarked before we are uncertain what causes this pattern. These findings shed some light. The combination of measurement error and censoring tends to shift cases with extreme values of the exogenous variable into another observation period, thus decreasing the correspondence between events and the X variables. Hence, we estimate a weaker association. Without measurement error censoring has the opposite effect, namely it censors extreme observations. Evidently, removal of extreme observations leads to the estimation of a stronger association between the exogenous variables and the rate. Frankly, we expected an opposite pattern because population variability is reduced.

If we turn again to Table 8, two additional remarks seem appropriate. First, while measurement error might shift the direction of the bias, it also increases its magnitude. Note that in most instances in Table 8 the

absolute value of the bias is greater when there is measurement error. These results would be discouraging if not for our second observation--despite this increase the %-bias remains low, always less than 12%. We regard these results as excellent for data with only 1.5 reliability.

Tables 9 and 10 provide further support for our assessment of MLE's performance under this condition. In Table 9 we see that measurement error slightly increases the variance of estimates for uncensored data. When censoring is present, however, estimator variance is actually decreased. Table 10 displays the mean-squared error. Note that MLE's performance when there is measurement error is comparable to its quality when there is no measurement error. In fact, at the highest censoring level, estimates are often better for the data with measurement error. These results follow from the decreased variance of the estimates when there is measurement error.

Measurement Error in the Exogenous Variables. Our final experiment also concerns measurement error, but in the exogenous (X) variables rather than the waiting time. The experiment is conducted as follows. Waiting times are generated as before, but the true exogenous variables are not used. Instead, a normal disturbance of mean zero and variance one-tenth is added to each exogenous variable. These altered exogenous variables are then used in conjunction with the waiting time variable generated by the true exogenous variables.

Results. Table 11 presents the results for this experiment when the correlation between the exogenous variables is zero. We find first that measurement error in the causal variables increases the magnitude of the bias but does not affect the directional pattern. Second, both the bias

and the percent-bias remain low despite only .9 reliability in the exogenous variables. Third, estimator variance is often lowered by the measurement error in the causal variables. Fourth, the mean-squared error shows that MLE yields robust estimates even with measurement error in the exogenous variables.

These results present a good contrast with those in which measurement error has created false censoring. When false censoring occurs, measurement error confounds and changes the quality of the estimates. Without false censoring, measurement error in both the time and exogenous variables seems to slightly increase the bias of estimates but leave the pattern intact. Figure 4 shows these different effects for a 60% level of censoring. In this figure, the curve for the case of causal variables measured with error does not contain cases with false censoring while the curve for time altered estimates curve does. Note that the curve with false censoring shifts towards zero and is more dispersed. In contrast, the curve of estimates for data with only measurement error in the exogenous variables closely resembles that for the data with no measurement error. These findings suggest that the major confounding problem of measurement error is false censoring.

5. Summary and Conclusion

We find these results concerning the performance of MLE in rate models to be encouraging. We have replicated and extended the results of earlier studies which show that MLE performs remarkably well. Further, we have seen that measurement error in the waiting time and in exogenous variables causes only minimal deterioration in estimator quality. In fact, our

studies strongly suggest that the most pronounced effects of measurement error can be avoided if the investigator does not impose arbitrary censoring points upon complete event-histories. When data on the timing of natural events are available, they should be used in entirety. Investigators following this advice can then place a high degree of confidence in maximum likelihood estimates of the time-dependent rate at which events occur.

FOOTNOTES

1. See Spilerman (1972), Sørensen (1975), Tuma (1976), Davis (1978) and Schuessler (1978) for recent examples.

2. A censored observation occurs when a unit under study does not experience an event during the observation period. For an extended discussion on alternative treatments of this situation, see Tuma and Hannan (1978).

3. The gamma-distributed deviate is generated with a rejection technique due to Johnk (1964) and developed by Phillips and Beightler (1972). The simulated distribution is

$$f(\gamma;\, a,b) = \frac{\gamma^{a-1}}{\Gamma(a)b^{a}} \exp\,(-\gamma/b)$$

where γ, a and b are always positive. For the random disturbance we use a = 10, b = .10 and for the time-dependent disturbance, $a = 10u_i$ and $b = .10u_i$. These parameters result in the stated means and variances.

4. For example, suppose censoring had been imposed at time = .5 and an individual's event occurred at time = .45. Normally this event would be observed. However, if the disturbance shifted the time to .52, then the case would be falsely censored. The phenomena can occur in the other direction as well, thus generating false events.

5. False censoring also can arise when measurement error is present and the observations are censored near the origin of a process. This occurs when event-histories are analyzed from a given data, i.e., after a person's twenty-first birthday. It should be noted, however, that false censoring is sometimes highly unlikely. In particular, when observations are censored at some point of intervention, false

censoring would require a respondent to report an historical event as occurring in the future or vice versa, both of which are highly improbable.

6. We denote duration in spell j for individual i as u_{ij} and waiting time until event j for individual i as t_{ij}.

7. Specifically, we use a gamma disturbance with mean 1 and variance $.10(5-t_{ij})$. The process is constrained so that the ordering of events remains constant. Consequently, the timing of an event with measurement error must be later than the timing of the most previous event (also with measurement error). For example, let $t_1=1$ and $t_2=2$ and the time of event one with measurement error (t_1^*) be 1.5. Then, t_2^* is constrained to be greater than 1.5 (the value of t_1^*). Thus, for $t_1<t_2<\cdots<t_j$ we must have $t_1^*<t_2^*<t_3^*<\cdots<t_j^*$. When waiting times do not satisfy this constraint, the disturbance is redrawn and t^* is recalculated.

8. This result might be explained by the constraint requiring the order of events to remain constant. For example, if an early waiting time for an individual is made dramatically longer by the disturbance, then subsequent events will be shifted toward the end of the process by the ordering constraint. While waiting times will more frequently be lowered, only one large increase in timing of events is required to cause events to "stack up" near the end of the process. Since the time-dependent disturbance gives a greater variance to earlier waiting times, this phenomena becomes more likely. Consequently, events are more likely to "stack up" in the fifth year and MLE will estimate an increased rate.

9. For example, when $\rho_{X_1X_2}=0$, for parameter -4, -1 and 1, the expected

time until a change of state is exp {5} and the expected rate of change is exp {-3}. When Fennell, Tuma and Hannan's (1977) parameter of -4, -2, and 2 are used, the expected time until change is exp {8} and the expected rate is 1. The differences between the variances are even greater.

10. When $\rho_{X_1X_2}=0$, the expected time until an event is exp {5} and the variance of waiting times is 2exp {12}-exp {10}. The disturbance changes the variance to 3exp {12}- $\frac{3}{2}$exp {10} + $\frac{1}{2}$exp {5} but does not change the expected value.

REFERENCES

Bartholomew, D.V.

1957 "A problem in life testing." Journal of the American Statistical Association 52: 350-56.

1959 "Note on the measurement and prediction of labour turnover." Journal of the Royal Statistical Society A, 122: 232-39.

1963 "The sampling distribution of an estimate arising in life testing." Technometrics 5: 361-74.

Breiman, L.

1969 Probability and Stochastic Processes. Boston: Houghton Mifflin.

Davis, J. A.

1978 "Studying categorical data over time." Social Science Research 7: 151-79.

Dhrymes, P.

1970 Econometrics: Statistical Foundations and Applications. New York: Harper and Row.

Fennell, M. L., N. B. Tuma and M. T. Hannan

1977 "Quality of maximum likelihood estimates of parameters in a log-linear rate model." Technical Report No. 59, Laboratory for Social Research, Stanford University.

Gill, P. E. and W. Murrary

1972 "Two methods for the solution of linearly constrained and unconstrained optimization problems." National Physical Laboratory Report NAC 25 (November).

Hannan, M. T., N. B. Tuma and L. P. Groeneveld

1977 "Income and marital events: evidence from an income maintenance experiment." American Journal of Sociology 82, 6 (May): 1186-1211.

Johnk, M. D.

1964 "Erzung von betaverteiler und gammaverteiler zufallzahlen." Metrika 8(2).

Keeley, M. C.

1975 "A Monte Carlo test of Nancy Tuma's maximum likelihood 'RATE' program." Stanford Research Institute, Menlo Park, Calif.

Mendenhall, W. and E. H. Lehman, Jr.

1960 "An approximation to the negative moments of the positive binomial useful in life testing." Technometrics 2: 227-42.

Phillips, D. T. and C. S. Beightler

1972 "Procedure for generating gamma variates with non-integer parameter sets." Journal of Statistical Computation and Simulation 1: 197-208.

Schuessler, K. (ed.)

1978 Sociological Methodology 1979. San Francisco: Jossey-Bass.

Sørensen, A.

1975 "The structure of intragenerational mobility." American Sociological Review 40: 456-71.

Spilerman, S.

1972 "The structure of intragenerational mobility." American Sociological Review 40: 174-200.

Tuma, N. B.

1976 "Rewards, resources and the rate of mobility: a nonstationary multivariate stochastic model." American Sociological Review 41: 338-60.

Tuma, N. B. and D. Crockford

1976 Invoking RATE. Stanford Research Institute, Menlo Park, Calif. Mimeo.

Tuma, N. B. and M. T. Hannan

1978 "Approaches to the censoring problem in analysis of event histories," in K. Schuessler (ed.), Sociological Methodology 1979. San Francisco: Jossey-Bass.

Tuma, N. B., M. T. Hannan and L. P. Groeneveld

1979 "Dynamic analysis of event histories." American Journal of Sociology 84. In press.

Table 1. Quality of a Maximum Likelihood Estimator of a Constant Rate Model with Simulated Measurement Error in the Waiting Time to an Event (α=1.0; No. of Samples=100).

Sample Size	Degree of Censoring	No Measurement Error		Random Multiplicative Error		Time-Dependent Multiplicative Error	
		Mean $\hat{\alpha}$	Variance $\hat{\alpha}$	Mean $\hat{\alpha}$	Variance $\hat{\alpha}$	Mean $\hat{\alpha}$	Variance $\hat{\alpha}$
25	0	1.082	.048	1.273	.078	1.364	.089
	.20	1.048	.075	1.319	.112	1.306	.116
	.50	1.040	.116	1.371	.172	1.330	.161
	.80	.973	.269	1.382	.422	1.115	.344
	.90	.999	.482	1.367	.751	1.051	.524
50	0	1.028	.023	1.227	.038	1.282	.048
	.20	1.021	.035	1.302	.043	1.341	.052
	.50	1.023	.052	1.320	.050	1.291	.065
	.80	.987	.092	1.338	.106	1.136	.110
	.90	1.003	.209	1.374	.272	1.037	.221
100	0	1.009	.010	1.201	.016	1.243	.024
	.20	1.016	.017	1.293	.027	1.324	.022
	.50	1.032	.023	1.325	.033	1.298	.029
	.80	1.013	.063	1.376	.081	1.139	.067
	.90	1.003	.109	1.358	.135	1.095	.116

Table 2. Comparison of Alternative Maximum Likelihood Estimators of a Multiple Event Constant Rate Model with Simulated Measurement Error in the Waiting Time to an Event (α=1.0; No. of Samples=100).

Estimator	Sample Size	No Measurement Error		Random Multiplicative Error		Time-Dependent Multiplicative Error	
		Mean	Variance	Mean	Variance	Mean	Variance
$\hat{\alpha}_1$ All-event MLE for first year	50	1.003	.029	1.221	.040	1.314	.034
	100	1.008	.012	1.241	.017	1.327	.016
$\hat{\alpha}_2$ Backward MLE for first year	50	1.012	.040	1.265	.055	1.280	.034
	100	--	--	--	--	--	--
$\hat{\alpha}_3$ All-event MLE for five years	50	.996	.004	.996	.004	.996	.004
	100	1.007	.002	1.007	.002	1.007	.002
$\hat{\alpha}_4$ All-event MLE for fifth year	50	1.016	.023	.715	.020	1.353	.043
	100	1.025	.012	.722	.009	1.392	.025
$\hat{\alpha}_5$ Backward MLE for fifth year	50	1.119	.001	1.078	.001	1.171	.002
	100	1.112	--	1.082	--	1.174	--
$\hat{\alpha}_6$ MLE for first event	50	1.012	.027	1.238	.044	1.353	.063
	100	1.019	.011	1.258	.017	1.382	.037
$\hat{\alpha}_7$ MLE for all events after the first	50	.998	.005	.955	.004	.940	.004
	100	1.006	.003	.962	.002	.947	.003

Table 3. Bias of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$		$\alpha_1 = 1$		$\alpha_2 = 2$	
		N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.034	.013	.003	-.014	-.005	-.007
ρ=0.0	60% Censored	-.001	-.033	.044	.010	.044	.045
	80% Censored	-.171	-.125	.076	.033	.167	.107
	Uncensored	.033	.018	-.026	.006	-.002	-.015
ρ=0.5	60% Censored	-.012	-.006	.038	-.009	.022	.031
	80% Censored	-.274	-.141	.097	.012	.163	.112
	Uncensored	.037	.012	.015	-.010	-.001	-.010
ρ=-0.5	60% Censored	-.006	-.034	.063	.013	.066	.050
	80% Censored	-.083	-.094	.034	.018	.099	.095
		$\alpha_0 = -4$		$\alpha_1 = -1$		$\alpha_2 = 1$	
	Uncensored	.034	.013	.003	-.014	-.005	-.007
ρ=0.0	60% Censored	.001	-.008	-.008	-.027	.001	.002
	80% Censored	-.133	-.085	-.078	-.066	.074	.045
	Uncensored	.034	.013	+.006	-.010	-.005	-.008
ρ=0.5	60% Censored	.003	-.013	-.003	-.027	.014	.008
	80% Censored	-.025	-.024	-.018	-.030	.029	.016
	Uncensored	.034	.013	-0-	-.017	-.005	-.008
ρ=-0.5	60% Censored	.001	-.012	-.018	-.031	.006	.007
	80% Censored	-.166	-.103	-.077	-.069	.098	.052

Table 4. Percent-Bias of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$		$\alpha_1 = 1$		$\alpha_2 = 2$	
		N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.9%	.3%	.3%	-1.4%	-.3%	-.4%
ρ=0.0	60% Censored	-0-	-.8%	4.4%	1.0%	2.2%	2.3%
	80% Censored	-4.3%	-3.1%	7.6%	3.3%	8.4%	5.4%
	Uncensored	.8%	.5%	-2.6%	.6%	-.1%	-.8%
ρ=0.5	60% Censored	-.3%	-.2%	3.8%	-.9%	1.1%	1.6%
	80% Censored	-6.9%	-3.5%	9.7%	1.2%	8.2%	5.6%
	Uncensored	.9%	.3%	1.5%	-1.0%	-0-	-.5%
ρ=-0.5	60% Censored	-.2%	-.9%	6.3%	1.3%	3.3%	2.5%
	80% Censored	-2.1%	-2.4%	3.4%	1.8%	5.0%	4.8%
		$\alpha_0 = -4$		$\alpha_1 = -1$		$\alpha_2 = 1$	
	Uncensored	.9%	.3%	.3%	-1.4%	-.5%	-.7%
ρ=0.0	60% Censored	-0-	-.2%	-.8%	-2.7%	.1%	.2%
	80% Censored	-3.3%	-2.1%	-7.8%	-6.6%	7.4%	4.5%
	Uncensored	.9%	.3%	.6%	-1.0%	-.5%	-.8%
ρ=0.5	60% Censored	.1%	-.3%	-.3%	-2.7%	1.4%	.8%
	80% Censored	-.6%	-.6%	-1.8%	-3.0%	2.9%	1.6%
	Uncensored	.9%	.3%	-0-	-1.7%	-.5%	-.8%
ρ=-0.5	60% Censored	-0-	-.3%	-1.8%	-3.1%	.6%	.7%
	80% Censored	-4.2%	-2.6%	-7.7%	-6.9%	9.8%	5.2%

Table 5. Variance of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$		$\alpha_1 = 1$		$\alpha_2 = 2$	
		N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.025	.011	.029	.013	.024	.013
ρ=0.0	60% Censored	.079	.042	.069	.028	.085	.041
	80% Censored	.490	.196	.162	.064	.340	.111
	Uncensored	.023	.011	.028	.017	.029	.018
ρ=0.5	60% Censored	.102	.051	.016	.048	.097	.054
	80% Censored	.730	.208	.177	.103	.339	.122
	Uncensored	.023	.010	.041	.021	.036	.018
ρ=-0.5	60% Censored	.073	.037	.086	.042	.120	.058
	80% Censored	.254	.097	.169	.079	.266	.096
		$\alpha_0 = -4$		$\alpha_1 = -1$		$\alpha_2 = 1$	
	Uncensored	.025	.011	.029	.013	.024	.013
ρ=0.0	60% Censored	.071	.033	.068	.027	.061	.026
	80% Censored	.228	.100	.133	.055	.151	.052
	Uncensored	.025	.011	.031	.014	.032	.017
ρ=0.5	60% Censored	.054	.029	.073	.036	.078	.035
	80% Censored	.097	.049	.120	.045	.124	.051
	Uncensored	.025	.011	.042	.020	.032	.017
ρ=-0.5	60% Censored	.065	.030	.081	.037	.072	.033
	80% Censored	.294	.124	.182	.074	.186	.079

Table 6. Mean Squared Error of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$		$\alpha_1 = 1$		$\alpha_2 = 2$	
		N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.026	.011	.029	.013	.024	.013
ρ=0.0	60% Censored	.079	.043	.071	.028	.087	.043
	80% Censored	.519	.212	.168	.065	.368	.122
	Uncensored	.024	.011	.029	.017	.029	.018
ρ=0.5	60% Censored	.102	.051	.062	.048	.097	.055
	80% Censored	.805	.228	.186	.103	.366	.135
	Uncensored	.024	.010	.041	.021	.036	.018
ρ=-0.5	60% Censored	.073	.038	.090	.042	.124	.061
	80% Censored	.261	.106	.170	.079	.276	.105
		$\alpha_0 = -4$		$\alpha_1 = -1$		$\alpha_2 = 1$	
	Uncensored	.026	.011	.029	.013	.024	.013
ρ=0.0	60% Censored	.071	.033	.068	.028	.061	.026
	80% Censored	.246	.107	.139	.059	.156	.054
	Uncensored	.026	.011	.031	.014	.032	.017
ρ=0.5	60% Censored	.54	.029	.073	.037	.078	.035
	80% Censored	.098	.059	.120	.046	.125	.051
	Uncensored	.026	.011	.042	.020	.032	.017
ρ=-0.5	60% Censored	.065	.030	.081	.038	.072	.033
	80% Censored	.322	.135	.188	.079	.196	.082

Table 7. Bias of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$				$\alpha_1 = -1$				$\alpha_2 = -1$			
		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error	
		N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.034	.013	.053	.022	.003	-.014	-.010	-.015	-.005	-.007	-.017	-.013
ρ=0.0	60% Censored	.001	-.008	.290	.269	-.008	-.027	.088	.089	.001	.002	-.105	-.099
	80% Censored	-.133	-.085	.342	.364	-.078	-.066	.060	.069	.074	.045	-.072	-.098
	Uncensored	.034	.013	.053	.022	.006	-.010	-0-	-.007	-.005	-.008	-.019	-.015
ρ=0.5	60% Censored	.003	-.013	.274	.251	-.003	-.027	.103	.095	.014	.008	-.111	-.111
	80% Censored	-.025	-.024	.345	.352	-.018	-.030	.115	.109	.029	.016	-.095	-.111
	Uncensored	.034	.013	.053	.022	-0-	-.017	-.019	-.022	-.005	-.008	-.019	-.014
ρ=-0.5	60% Censored	.001	-.012	.279	.264	-.018	-.031	.039	.053	.006	.007	-.114	-.100
	80% Censored	-.166	-.103	.320	.374	-.077	-.069	.025	.054	.098	.052	-.059	-.088

Table 8. Percent-Bias of Estimates: Log-Linear Rate Model with Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$				$\alpha_1 = -1$				$\alpha_2 = -1$			
		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error	
		N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.9%	.3%	1.3%	.6%	.3%	-1.4%	-1.0%	-1.5%	-.5%	-.7%	-1.7%	-1.3%
ρ=0.0	60% Censored	-0-	-.2%	7.3%	6.7%	-.8%	-2.7%	8.8%	8.9%	.1%	.2%	-10.5%	-9.9%
	80% Censored	-3.3%	-2.1%	8.6%	9.1%	-7.8%	-6.6%	6.0%	6.9%	7.4%	4.5%	-7.2%	-9.8%
	Uncensored	.9%	.3%	1.3%	.6%	.6%	-1.0%	-0-	-.7%	-.5%	-.8%	-1.9%	-1.5%
ρ=0.5	60% Censored	.1%	-.3%	6.9%	6.3%	-.3%	-2.7%	10.3%	9.5%	1.4%	.8%	-11.1%	-11.1%
	80% Censored	-.6%	-.6%	8.6%	8.8%	-1.8%	-3.0%	11.5%	10.9%	2.9%	1.6%	-9.5%	-11.1%
	Uncensored	.9%	.3%	1.3%	.6%	-0-	-1.7%	-1.9%	-2.2%	-.5%	-.8%	-1.9%	-1.4%
ρ=-0.5	60% Censored	-0-	-.3%	7.0%	6.6%	-1.8%	-3.1%	3.9%	5.3%	.6%	.7%	-11.4%	-10.0%
	80% Censored	-4.2%	-2.6%	8.0%	9.4%	-7.7%	-6.9%	2.5%	5.4%	9.8%	5.2%	-5.9%	-8.8%

Table 9. Variance of Estimates: Log-Linear Rate Model with Normally-Distributed Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$				$\alpha_1 = -1$				$\alpha_2 = -1$			
		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error	
		N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.025	.011	.046	.027	.029	.013	.047	.021	.024	.013	.038	.027
ρ=0.0	60% Censored	.071	.033	.049	.027	.068	.027	.056	.029	.061	.026	.060	.034
	80% Censored	.228	.100	.098	.057	.133	.055	.085	.039	.151	.152	.104	.054
	Uncensored	.025	.011	.046	.027	.031	.014	.057	.025	.032	.017	.051	.036
ρ=0.5	60% Censored	.054	.029	.042	.027	.073	.036	.068	.030	.078	.035	.077	.043
	80% Censored	.097	.049	.066	.035	.120	.045	.091	.038	.124	.051	.114	.056
	Uncensored	.025	.011	.046	.027	.042	.020	.063	.035	.032	.017	.051	.036
ρ=-0.5	60% Censored	.065	.030	.062	.028	.081	.037	.082	.042	.072	.033	.086	.047
	80% Censored	.294	.124	.191	.084	.182	.074	.131	.064	.186	.079	.167	.076

Table 10. Mean Squared Error of Estimates: Log-Linear Rate Model with Normally-Distributed Variables. (All entries based on 100 observations.)

		$\alpha_0 = -4$				$\alpha_1 = -1$				$\alpha_2 = -1$			
		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error		No Measurement Error		Random Multiplicative Error	
		N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100
	Uncensored	.026	.011	.049	.027	.029	.013	.047	.021	.024	.013	.038	.027
ρ=0.0	60% Censored	.071	.033	.133	.099	.068	.028	.064	.037	.061	.026	.071	.044
	80% Censored	.246	.107	.215	.189	.139	.059	.089	.044	.156	.054	.109	.064
	Uncensored	.026	.011	.049	.027	.031	.014	.057	.025	.032	.017	.051	.036
ρ=0.5	60% Censored	.054	.029	.117	.090	.073	.037	.079	.039	.078	.035	.089	.055
	80% Censored	.090	.050	.188	.159	.120	.046	.104	.050	.125	.051	.123	.068
	Uncensored	.026	.011	.049	.027	.042	.020	.063	.035	.032	.017	.051	.030
ρ=-0.5	60% Censored	.065	.030	.140	.098	.081	.038	.084	.045	.072	.033	.099	.057
	80% Censored	.322	.135	.293	.224	.188	.079	.132	.067	.196	.082	.170	.084

Table 11. Quality of Maximum Likelihood Estimator for a Multivariate Log-Linear Rate Model with Simulated Measurement Error in the Normally-Distributed Exogenous Variables. (All entries based on 100 observations.)

	$\rho = 0.0$	$\alpha_0 = -4$				$\alpha_1 = -1$				$\alpha_2 = -1$			
		No Measurement Error		Random Additive Error in Exogenous Variables		No Measurement Error		Random Additive Error in Exogenous Variables		No Measurement Error		Random Additive Error in Exogenous Variables	
		N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100	N=50	N=100
Bias of Estimates	Uncensored	.034	.013	.037	.017	.003	-.014	-.029	-.016	-.005	-.007	-.007	-.003
	60% Censored	.001	-.008	.002	.004	-.008	-.027	-.075	-.045	.001	.002	.036	.007
	80% Censored	-.133	-.085	-.108	-.020	-.078	-.066	-.125	-.049	.074	.045	.065	.002
Percent-Bias of Estimates	Uncensored	.9%	.3%	.9%	.4%	.3%	-1.4%	-2.9%	-1.6%	-.5%	-.7%	-.7%	-.3%
	60% Censored	-0-	-.2%	.1%	.1%	-.8%	-2.7%	-7.5%	-4.5%	.1%	.2%	3.6%	.7%
	80% Censored	-3.3%	-2.1%	-2.7%	-.5%	-7.8%	-6.6%	-12.5%	-4.9%	7.4%	4.5%	6.5%	.2%
Variance of Estimate	Uncensored	.025	.011	.023	.013	.029	.013	.028	.012	.024	.013	.016	.006
	60% Censored	.071	.033	.061	.034	.068	.027	.062	.030	.061	.026	.067	.025
	80% Censored	.228	.100	.220	.113	.133	.055	.161	.060	.151	.052	.115	.046
Mean Squared Error of Estimates	Uncensored	.026	.011	.024	.013	.029	.013	.029	.012	.024	.013	.016	.006
	60% Censored	.071	.033	.061	.034	.068	.028	.068	.032	.061	.026	.068	.025
	80% Censored	.246	.107	.231	.113	.139	.059	.177	.062	.156	.054	.119	.046

FIGURE 1

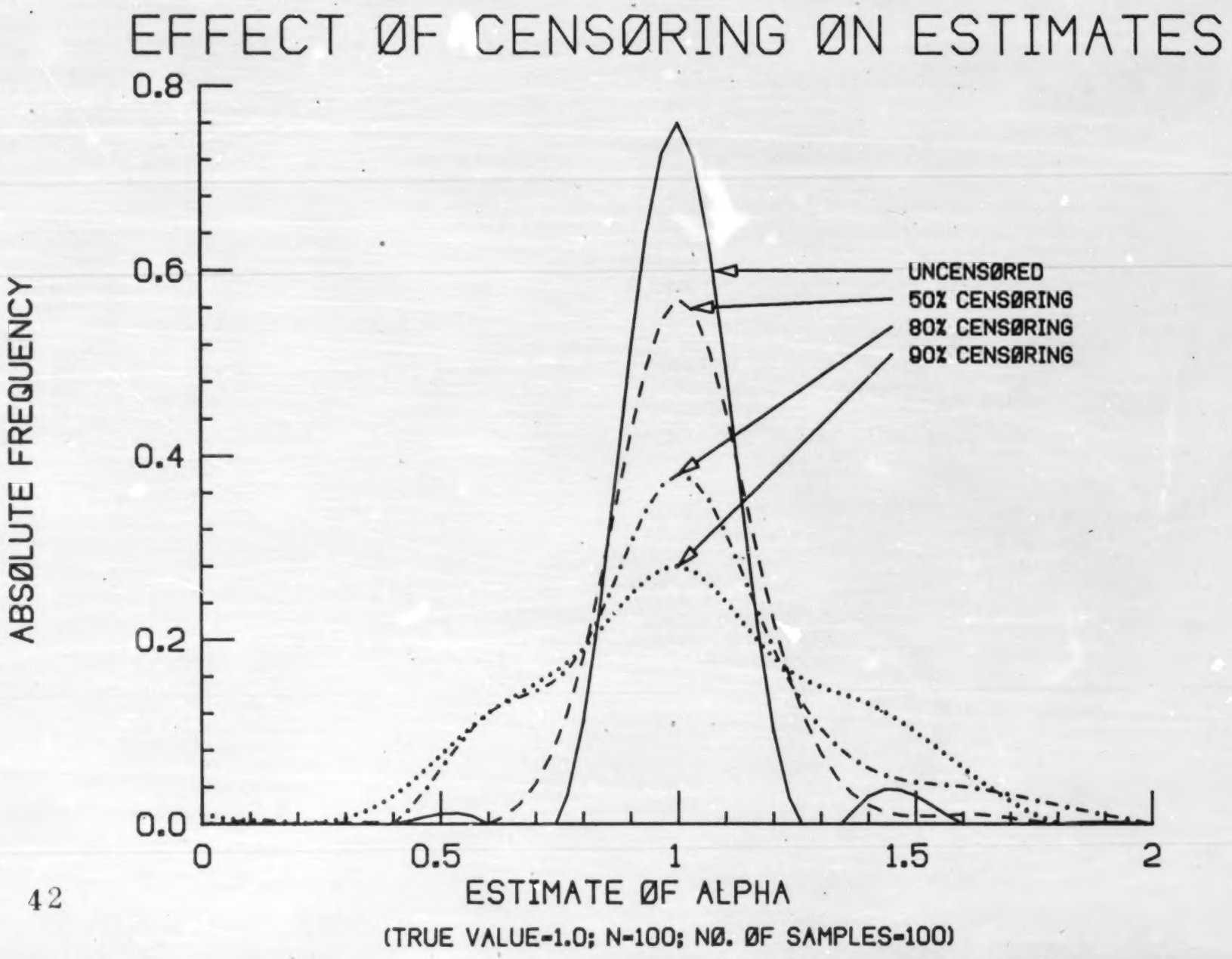
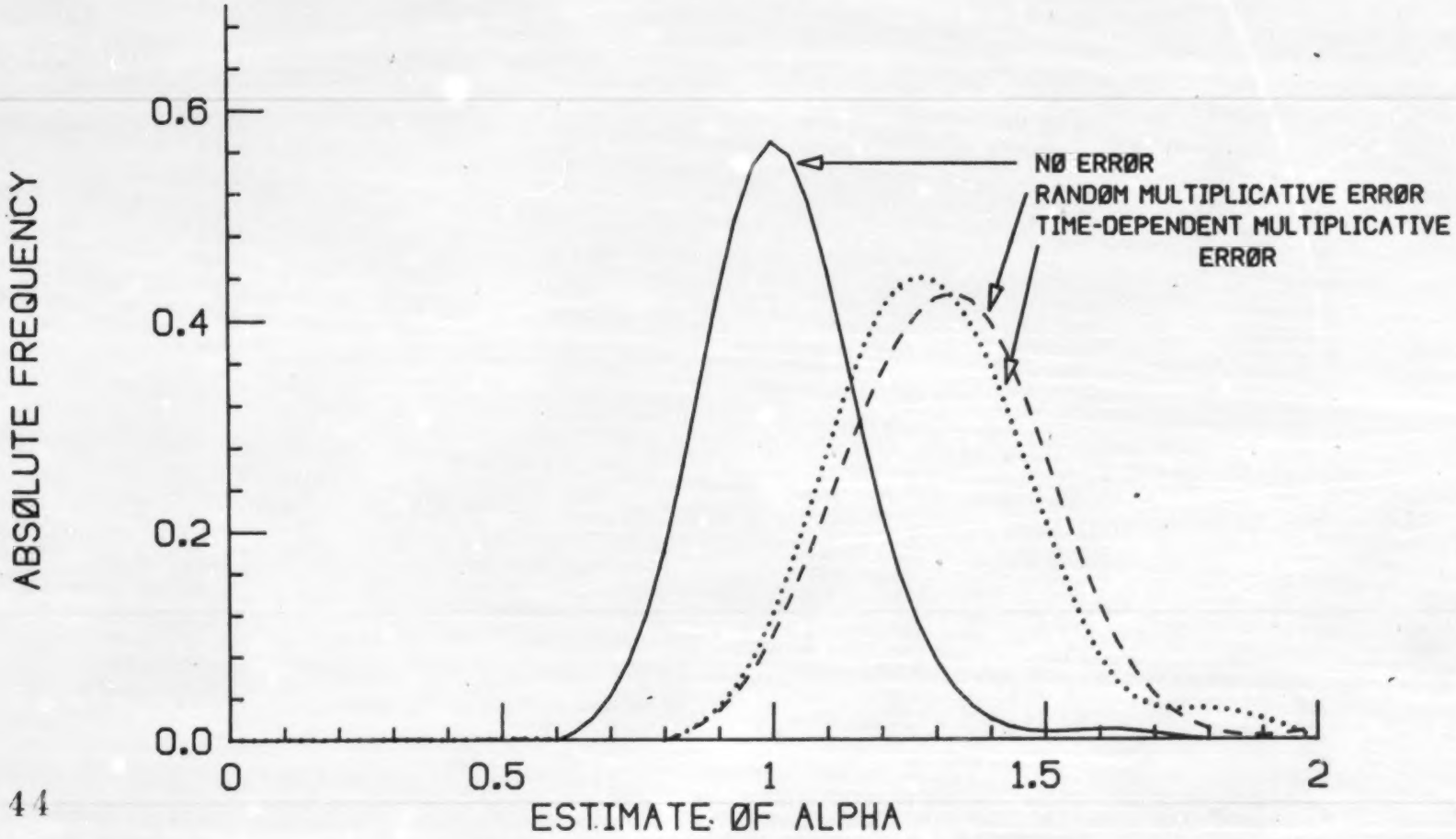

EFFECT ØF CENSØRING ØN ESTIMATES
ABSØLUTE FREQUENCY
0.8
0.6
0.4
0.2
0.0
0
0.5
1
1.5
2
ESTIMATE ØF ALPHA
UNCENSØRED
50% CENSØRING
80% CENSØRING
90% CENSØRING
(TRUE VALUE=1.0; N=100; NØ. ØF SAMPLES=100)

FIGURE 2

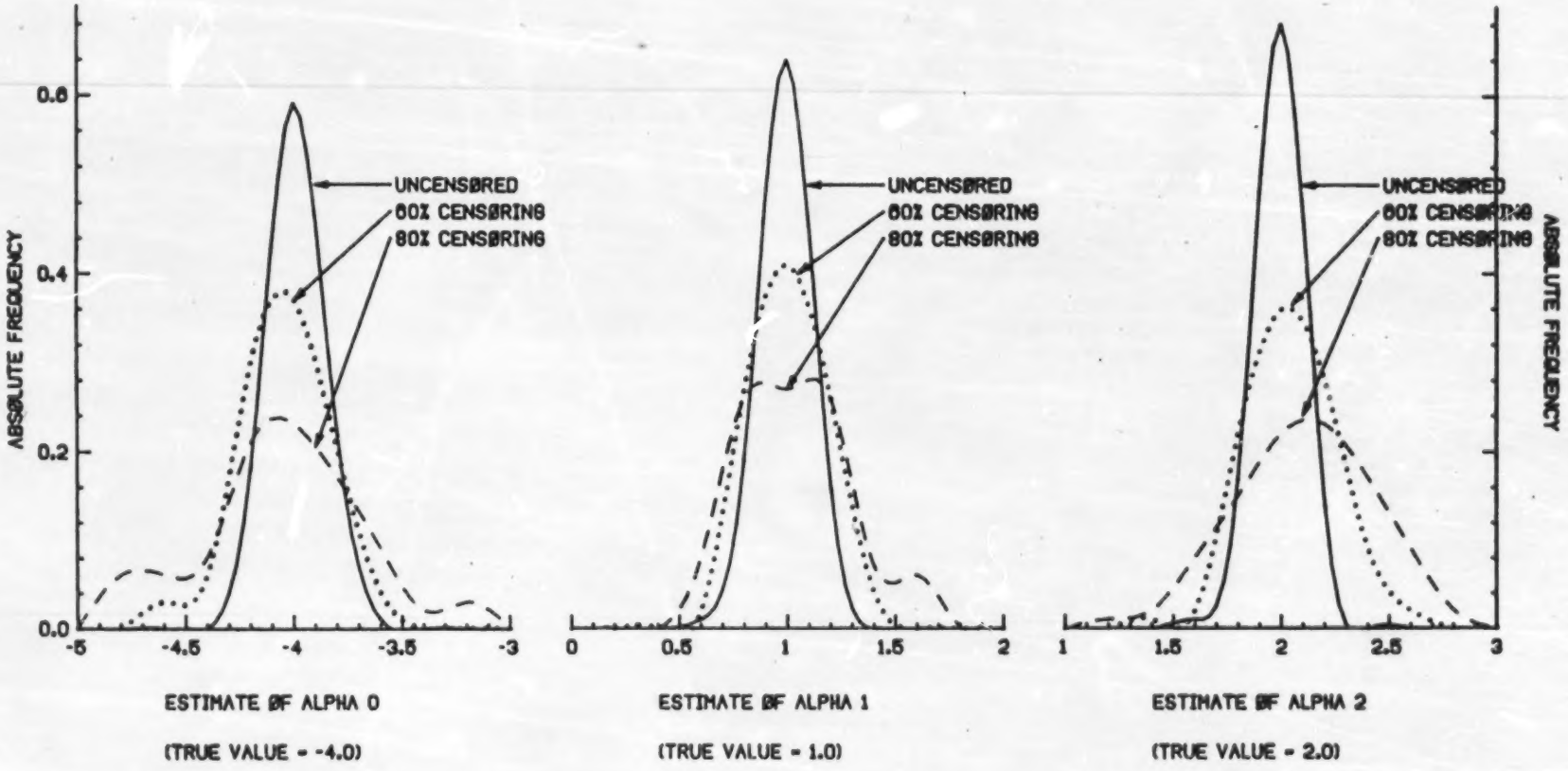
EFFECT ØF SIMULATED ERRØRS ØN ESTIMATES ØF 50% CENSØRED DATA



(TRUE VALUE-1.0; N-100; NØ. ØF SAMPLES-100)

FIGURE 3

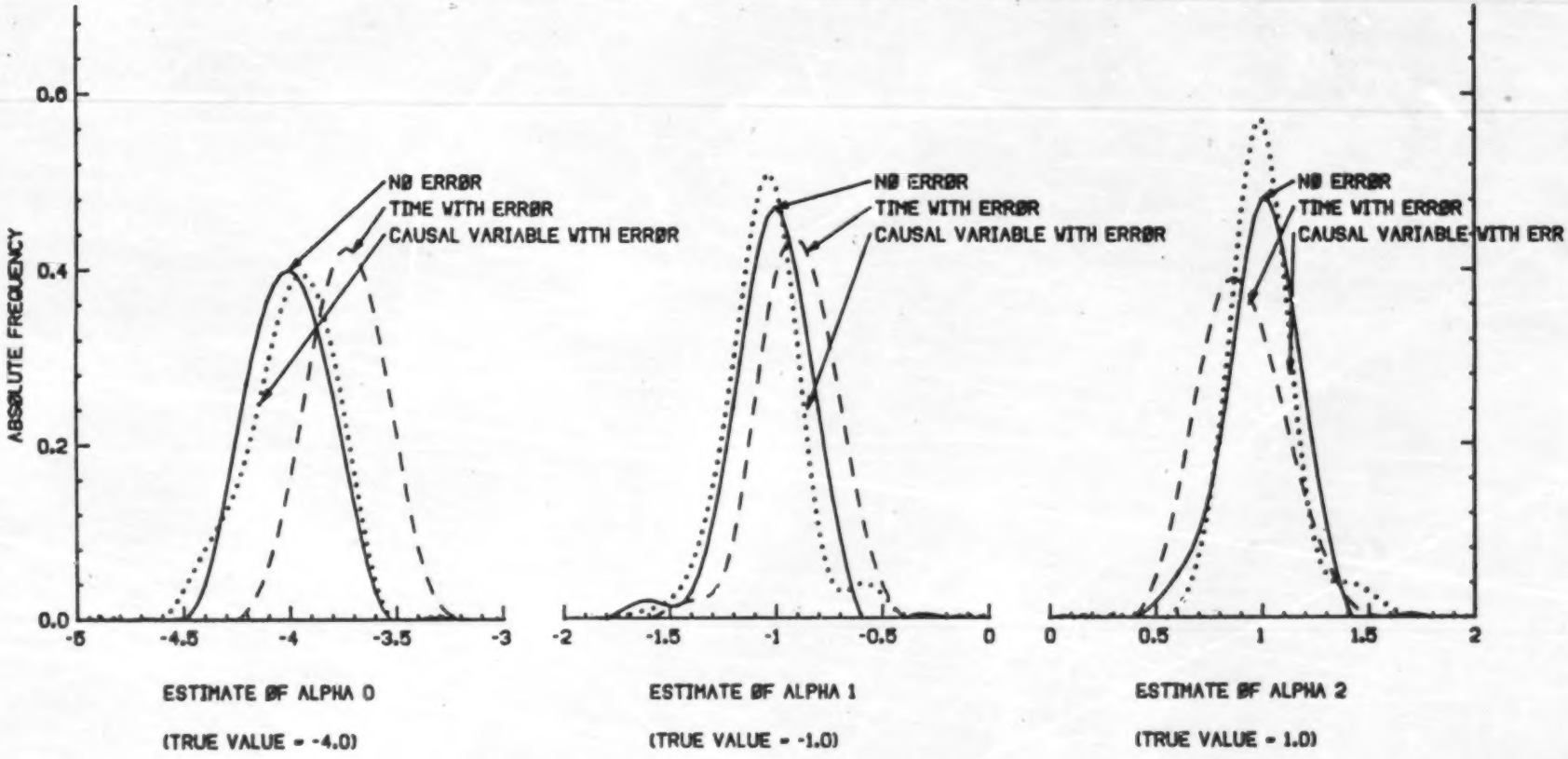
EFFECT ØF CENSØRING ØN ESTIMATES



RHØ=0.0; N=100; NØ. ØF SAMPLES=100

FIGURE 4

EFFECT ØF SIMULATED ERRØRS ØN ESTIMATES ØF 60% CENSØRED DATA



RHØ=0.0; N=100; NØ. ØF SAMPLES=100